AIAA AVIATION Forum
June 25-29, 2018, Atlanta, Georgia
2018 Aviation Technology, Integration, and Operations Conference

Field Testing of Vision-Based Surveillance System for Ramp Area Operations

Hui-Ling Lu [*], Jason Kwan [†], Andrew Fong [‡], and Victor H. L. Cheng [§]
Optimal Synthesis Inc., Los Altos, CA, 94022

The objective of this research is to develop a Vision-BAsed Surveillance System (VBASS) for ramp areas in airports. Conventional surveillance systems such as RADAR, Automated Dependent Surveillance - Broadcast (ADS-B), and Airport Surface Detection Equipment-Model X (ASDE-X) cater to taxiways and runways. Without suffering from multi-path errors, vision-based surveillance technologies are based on surveillance cameras located at certain vantage points in the ramp area. To illustrate the functionality of the VBASS system in a busy airport, we installed six cameras at the Charlotte (CLT) Douglas International Airport and demonstrated that VBASS is capable of tracking aircrafts under different lighting conditions in real time using the streaming videos obtained from the CLT cameras.

I. Introduction

The National Aeronautics and Space Administration (NASA) and Federal Aviation Administration (FAA) have been involved in extensive efforts to develop advanced concepts, technologies, and procedures for the Next Generation Air Transportation System (NextGen) [1]-[5]. NextGen seeks improvements in safety, efficiency, controller workload, and the capacity of the National Airspace System (NAS). The aforementioned improvements are in turn sought to be achieved by improvements in communications, surveillance, navigation, and automation systems. The objective of the current work is to develop a Vision-BAsed Surveillance System (VBASS) for ramp areas in airports.

Conventional surveillance systems such as RADAR, Automated Dependent Surveillance - Broadcast (ADS-B), and Airport Surface Detection Equipment-Model X (ASDE-X) cater to taxiways and runways. Designing a surveillance system for the ramp area poses the following challenges:

- Operation of RADAR transmitters is restricted in the ramp area due to safety concerns [6].
- The accuracy of conventional surveillance systems such as primary and secondary RADAR and Automatic Dependent Surveillance – Broadcast (ADS-B) deteriorates in the ramp area due to multi-path errors [7]. Figure 1 shows the RADAR dead zone at CLT. Figure 2 shows a sample surveillance error observed from real data at Dallas Fort Worth (DFW). The big jump in the blue line (aircraft track) indicates large errors close to the gate.
- Most of the above surveillance systems not only require aircraft equipage but also require the aircraft avionics to be powered 'On' for enabling the surveillance system. Typically aircraft power 'Off' their avionics during passenger and fuel loading.

Vision-based surveillance technologies are based on surveillance cameras located at certain vantage points in the ramp area. The cameras required for such a surveillance system are commercially available [8], [9]. Whereas cameras operating in visible wavelength are sufficient for clear weather daytime operations, infrared cameras [10], [11] are necessary for nighttime and all weather conditions such as rain, snow, and fog. Both the visible wavelength and the infrared cameras are currently available as commercial-off-the-shelf components [8]-[11]. Some of the airports such as Dallas/Fort Worth International Airport (DFW), San Francisco International Airport (SFO), and Charlotte Douglas International Airport (CLT) already have cameras installed in their ramp areas.

[*] Senior Research Scientist, Optimal Synthesis Inc., Los Altos, CA 94022.
[†] Research Engineer, Optimal Synthesis Inc., Los Altos, CA 94022.
[‡] Research Engineer, Optimal Synthesis Inc., Los Altos, CA 94022.
[§] President, Optimal Synthesis Inc., Los Altos, CA 94022, AIAA Associate Fellow.

Figure 1. RADAR Dead Zone at CLT (Shown Using the Red Polygon)

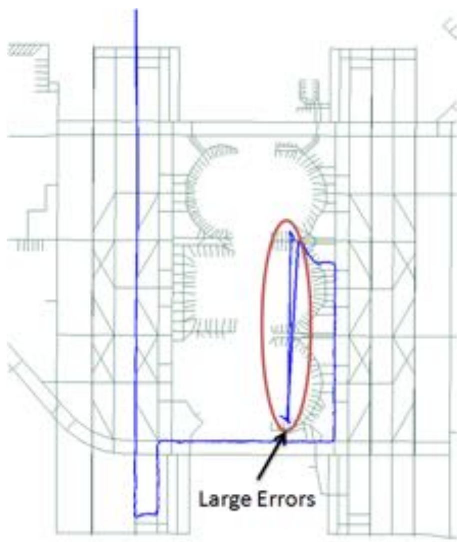

Figure 2. Surveillance Errors in Ramp Area at DFW

II. VBASS Overview

VBASS takes in as input the video feed from the cameras and processes it using computer-vision algorithms. VBASS could approximately update at 1Hz the following surveillance data:

- Total number of aircraft in the ramp area
- Current 3D location of all the aircraft in an inertial frame of reference with additional localization information such as gate and ramp spot
- Current speed and heading
- Aircraft type and airline for all the aircraft
- Pushback status of departure flights from the gate

Figure 3 shows a functional block diagram of VBASS which consists of the following algorithmic modules: (i) Aircraft & Ground-Vehicle Detection Module, (ii) Aircraft Localization Module, and (iii) Data Fusion Module. The purpose of the Detection module is to detect the presence of aircraft or a ground-vehicle in each video frame obtained from all the cameras. The Localization module assigns a 3D location to each identified aircraft. The Data-Fusion module correlates past and present aircraft detection/location data with flight plans and other information to maintain continuous and accurate aircraft tracks.

VBASS is expected to aid in both efficiency improvements in ramp area operations and safety improvements in overall airport surface operations. VBASS can be used in real-time mode or it can be used in post-processing (playback) mode. In real-time mode VBASS can aid planners such as the Spot and Runway Departure Advisor (SARDA) [12] by providing real-time updates on the pushback status of the aircraft; VBASS can also help monitor the safety of ramp-area operations in an automated manner. In post-processing mode VBASS can help identify from video surveillance data, patterns of interest in ramp area operations such as the paths from the gates to spot, and the transit time from the gate to the ramp spot.

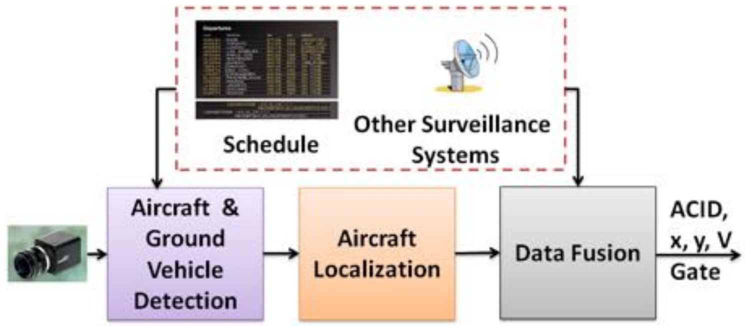

Figure 3. VBASS Functional Block Diagram

We installed six cameras monitoring the traffic in terminal B-C alley. These cameras have been adjusted to cover as much of the view of the ramp area as possible. Two cameras were installed on the ramp tower overlooking the B-C alley. Two cameras were installed on the roof of the C concourse and another two cameras were installed on the roof of the B concourse. Since two cameras at the terminal C cannot be installed at a higher altitude, the coverage area at the B terminal side does not include the first gate. Figure 4 shows the viewing coverage area from cameras installed at the ramp tower and the C concourse.

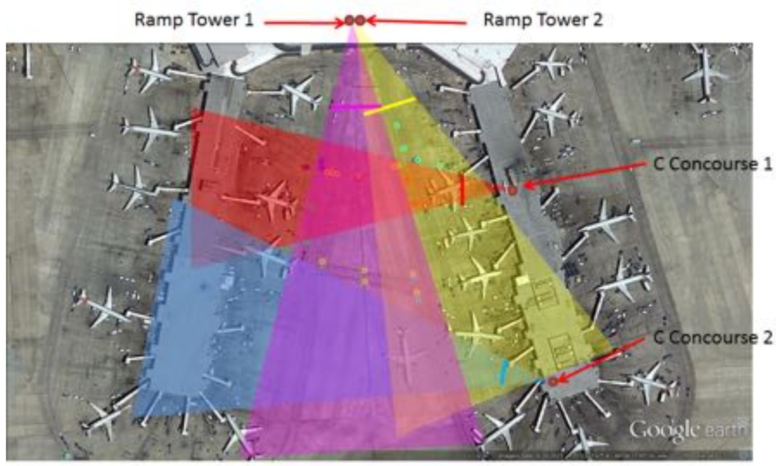

Figure 4. Camera Coverage Area for Cameras Located at the Ramp Tower and C Concourse

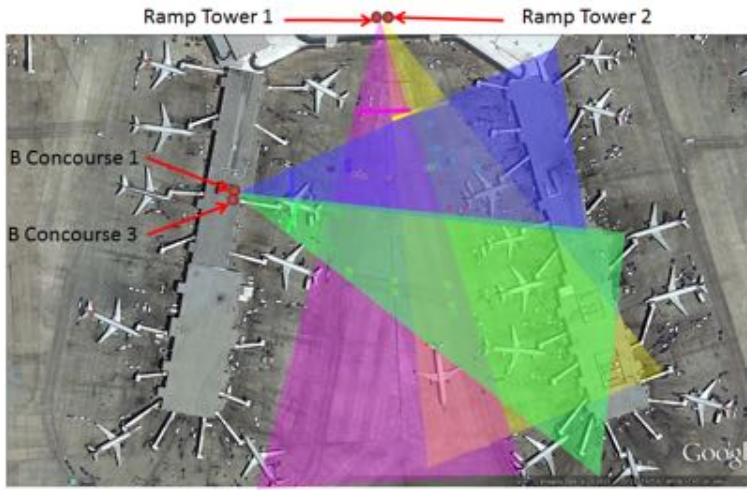

Figure 5. Camera Coverage Area for Cameras Located at the Ramp Tower and B Concourse

Figure 5 shows the viewing coverage area from cameras installed at the ramp tower and the B concourse. These cameras cover most of the ramp area between B and C concourses except at the very end of both concourses and the first gate of the B concourse. Figure 6 shows sample views from six cameras installed at CLT airport.

Figure 6. Sample Views of the CLT Cameras

III. VBASS Implementation

Localization and tracking of aircrafts were demonstrated using the video clips recorded at ramp area of the B-C concourse of the CLT airport. Section A describes the VBASS software architecture. Section B briefs the aircraft detection implementation. Section C describes the aircraft localization module and Section D illustrates the data fusion module that combines detection results from all cameras along with integration of the flight schedules.

A. Software Architecture

The overall system architecture is shown in Figure 7. The software is designed as a set of several modules that communicate with one another using socket connections. Each module runs as its own process with the Data Fusion module serving as the master process. The Data Fusion module receives detected object input from each of the detector modules and sends data to the top-down map display (See Figure 8) and aircraft status dashboard modules (See Figure 9).

To integrate the resulting aircraft surveillance data with other traffic information, VBASS has combined the available traffic data from the System Wide Information Management (SWIM) [13] data feed with the vision-based surveillance results and shown the aggregate traffic information on the top-down display. In our current implementation, we modified the Asset Player software originally developed by NASA as the top-down display that shows the aircraft movement at the ramp area.

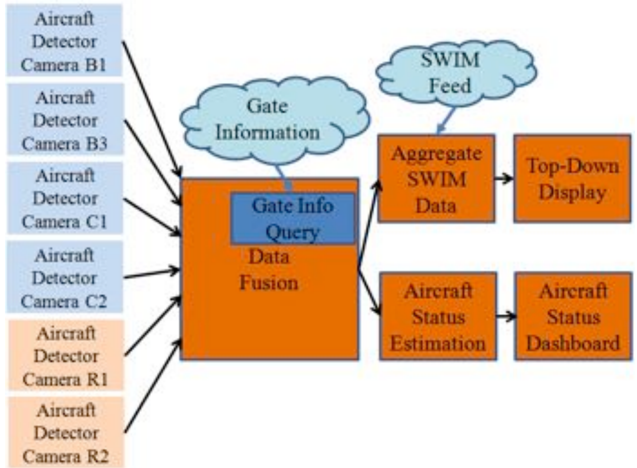

Figure 7. VBASS Software Architecture

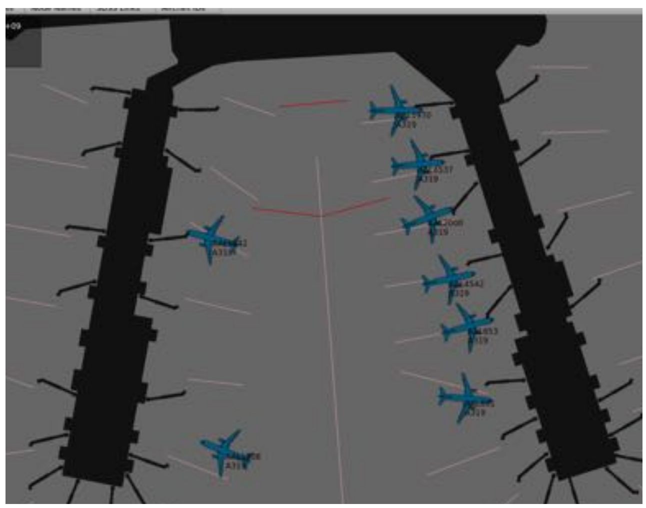

Figure 8. Top-Down Display

Callsign	Serial No.	Gate	Category	Status	Latitude	Longitude
AAL1808	4204951491	B12	Departure	Pushback	35.217950	-80.94451
AAL1842	4204952009	B6	Arrival	Gate	35.219115	-80.94460
AAL1969	4204969439	C11	Departure	Gate	35.218635	-80.94289
AAL1970	4204969561	C3	Arrival	Gate	35.219817	-80.94337
AAL2000	4207062286	C7	Departure	Gate	35.219225	-80.94316
AAL4470	4211628029	C5	Departure	Gate	35.219520	-80.94322
AAL4542	4211644799	C9	Arrival	Gate	35.218894	-80.94302
AAL601	1540340079	C13	Departure	Gate	35.218241	-80.94291

Figure 9. Aircraft Status Dashboard (Partial View)

B. Aircraft Detection

Computer-vision based object detection is based on the "Discriminatively Trained Deformable Part Models" approach originally proposed by Felzenszwalb [14], [15]. This approach is a supervised-learning approach. Pre-trained deformable part models (DPM) are required for detecting objects of interests. Using video data collected at the CLT airport, we partitioned the database to train different aircraft DPM models such as daytime aircraft models and nighttime aircraft models. In addition to aircraft DPM models, to support the aircraft pushback detection task, we also trained aircraft tail models using videos collected from one of the ramp camera that focus on the side view of the C terminal.

Specifically designed for spotting the tail pieces using ramp tower camera that focuses at the C terminal side, a simple one-component DPM model turns out to be sufficient to detect tails of the aircraft parking at the C gates.

Since aircrafts parking at the C terminal side have similar orientations and the aircraft tails are rigid with simple shapes, one-component DPM model without parts is used to reduce detection complexity. Our testing has shown effective detections using the simplified one-component root model. Figure 10 shows the daytime DPM aircraft tail model. The left column shows the root filter, the middle column shows the models for the parts, and the right column shows the deformation cost. The shape of the tail can be seen vividly from the root filter. Since the model has left-right symmetry by design, one–component model includes two mirrored orientations. Figure 11 shows one detection example using the image taken from the ramp tower camera 1. All side-facing aircraft tails in view are detected correctly for this testing video clip with adjusted detection threshold using set-aside validation data.

In addition to aircraft detection at the ramp area, to disambiguate the clutter view of the aircrafts from one of the ramp camera, we also built concourse aircraft detector modules that use Support Vector Machine (SVM) binary classifiers to determine if an aircraft is parked at a particular gate. A gate region in the field of view of the camera can be pre-specified. In each gate region of the video snapshot, SVM binary classifier model and bag-of-words dictionary are learnt in order to perform binary classification of the aircraft presence. For each gate region in the image, if the SVM classifier determines that an aircraft is parked at the gate then a box is drawn around the gate region in the image to indicate a positive detection and the gate's bounding box is added to the list of data to send to the data fusion module for the current timestamp.

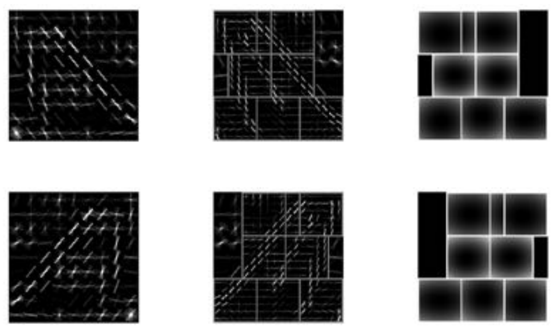

Figure 10. Aircraft Tail DPM Model

Figure 11. Example Tail Detection Result

The concourse detector module provides a video display window to show the status of the SVM classifier at each gate. If a flight is detected at a gate, then a box is drawn around the gate region. An example of a concourse aircraft detector video display is shown in Figure 12.

C. Aircraft Localization via Homography Mapping

Aircraft localization is performed by combing the aircraft detection results from multiple cameras after converting the detected locations in the image plane to the common top-down ground plane via the homography mapping transformation.

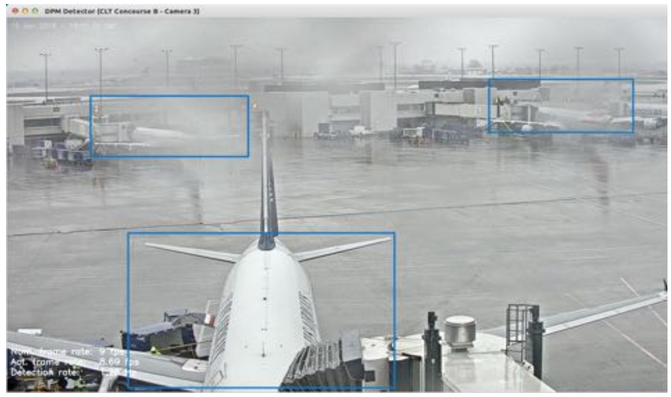

Figure 12. Aircraft Detection at the Gate via the SVM Classifier

Under the assumption of the pinhole camera model, any two images of the same planar surface can be related by a 2D homography projective matrix. That is, we can find a linear transformation between the coordinates on a plane (such as the ground) and the image plane pixel location. The transformation once calibrated can be used to transition from image plane to ground plane and vice versa. The following equation demonstrates the homography transformation:

$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = H_{3\times3} \begin{bmatrix} X \\ Y \\ s \end{bmatrix} \tag{1}$$

where u, v are the homogeneous pixel plane coordinates to be scaled using w; and X, Y are the ground plane homogeneous coordinates to be scaled by s. In this example the ground plane is treated as $Z = constant$. The homography matrix H is an invertible matrix that can be used for conversions between the two coordinate systems. The H matrix is obtained by careful calibration, in which the user selects visually distinguishable points in the image plane and associates with them pre-determined world coordinates. The H matrix is then solved using numerical techniques after obtaining at least four such non-collinear correspondences.

Homography is a useful way for determining the world coordinates of certain image plane pixels using only a single camera. In order to find the mapping between each camera view and the top-down map display, we select visible landmarks on the ground from images of each surveillance camera and then correspond these landmarks to the latitude longitude locations using the Google map. For example, the image-plane key points for Ramp Camera 1 are shown in Figure 13 and the corresponding Google map plane coordinates are shown in Figure 14. Homography matrix for each camera view is computed from these correspondence pairs.

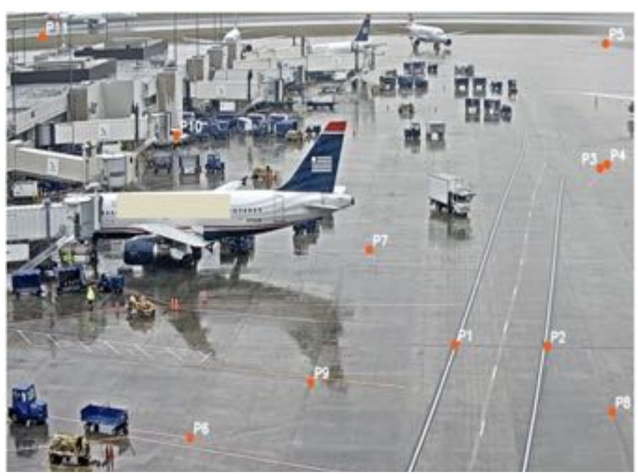

Figure 13. Image-Plane Key Points (Ramp 1 Camera)

The fitting error of the transformation is computed by applying the homography matrix to the image-plane key points and computing the error in the resulting latitude/longitude position in meters.

Table 1 shows the minimum, maximum, and average error for the homography mappings of training key points. Based on the current viewing angle, the homography mapping is very sensitive in the y-direction of the image plane. Small variations of only a few pixels in magnitude far from the camera (near the top of the image) can produce large position errors of tens of feet.

Figure 14. Google-Map Key Points (Ramp 1 Camera)

Table 1. Homography Mapping Errors

	Ramp Camera 1	Ramp Camera 2
Minimum Error [m]	0	0
Maximum Error [m]	1.4512	2.2339
Average Error [m]	0.5474	0.8803

D. Data Fusion

The Data-Fusion module correlates past and present aircraft detection/location data with flight plans and other information to maintain continuous and accurate aircraft tracks. The data fusion module serves as the master process for the VBASS software system; each of the other client processes (detectors, SWIM data, and aircraft status detector) connects to this module using TCP sockets. The data fusion module receives detector data from each of the connected detector modules and aggregates the data. The aggregated data is then sent to the SWIM data module and the aircraft status detector module.

When aggregating the data, the data fusion module collects the detected flight positions from each of the detector modules and merges the flight positions from different detectors, which correspond to the same flight into a single point.

Since flight schedule data provides prior information of gate occupancy and additional flight information such as flight ID, fusion of the VBASS detection results with the flight schedule data is also performed at the data fusion stage, which allows the VBASS system to associate the aircraft tracks with its flight ID and to enhance its performance by integrating the prior information.

IV. Results

The objective of this research is to demonstrate that VBASS can process video streams in real-time to detect aircraft in the ramp area and to provide aircraft positions on a top-down map display and other aircraft status information.

To ensure that the VBASS system performs under various lighting conditions, testing videos have been selected to cover a wide range of lighting and weather conditions. With input frame rate of 9 Hz, VBASS is capable of outputting surveillance results in 1Hz. With the adjustment of model choice and system parameter modifications for different lighting conditions, in general, VBASS can effectively detect and track aircrafts when the tracking target does not go out of the camera view or compromised with a very low signal-to-noise ratio for example due to rain drops on the camera cover.

Figure 15 shows one snapshot of the VBASS in action. The left-hand side shows camera view from Ramp 1 camera, Ramp 2 camera, and concourse B1 camera overlooking the front of the concourse C area. The right-hand sides of these figures show the detected aircraft on the top-down display and also an aircraft status GUI that tabulates the status of the aircraft tracks. Since the aircraft type is unknown in this example, all aircrafts are shown as A319 type. Figure 15 shows one aircraft taxiing out from the gate C7 on a rainy night. One may observe that the aircraft status for the gates farther away from the ramp tower such as the gate C15 may not be accurate due to larger homography mapping errors. We plan to incorporate gate location information at the data fusion stage in the future to reduce the impact from homography mapping errors.

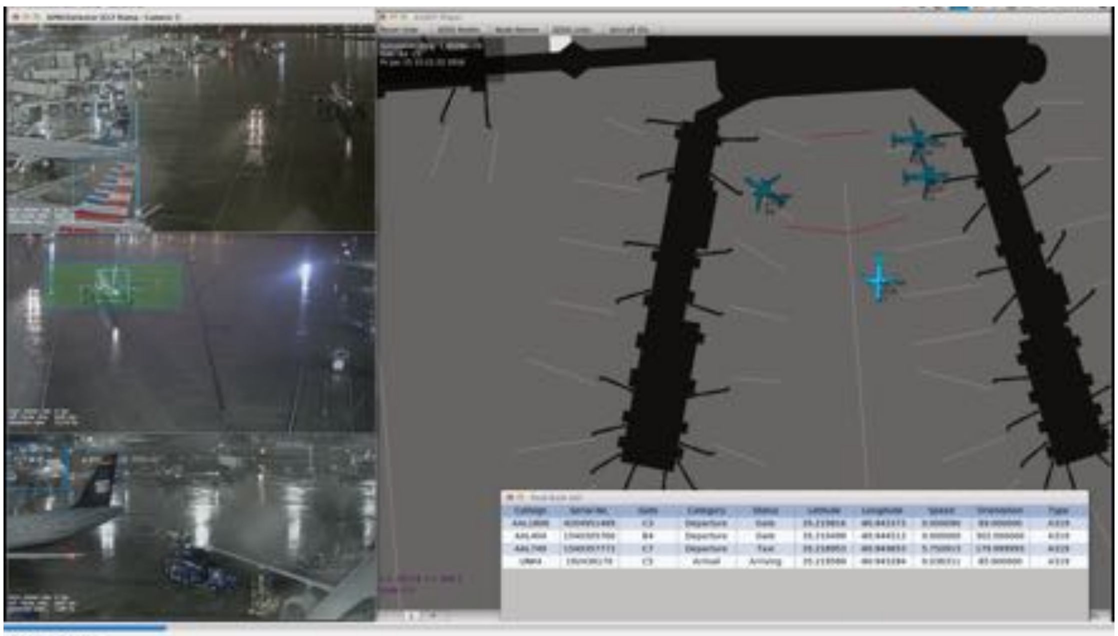

Figure 15. VBASS Result Snapshot (Aircraft Departure, Rainy Day, Night Time)

To obtain the VBASS baseline aircraft tracking performance, we conducted regression tests on VBASS using twenty testing video recordings that cover various lighting and weather conditions. The performance metrics include average aircraft tracking errors and average tracking dropout rate among all aircraft tracks. The average aircraft models trained by using recording data from 6am to 9pm are used to deduce the statistics of the aircraft tracking errors. Figure 16 illustrates the average root mean square (RMS) errors in the image plane in pixels for x-axis and for y-axis over all tracks in the testing scenarios. Figure 17 depicts the average RMS errors in feet over all tracks in the testing scenarios. Despite small errors in the image plane, the errors in the top-down plane could be large due to inaccurate mapping of ground projection points or homography mapping errors. Additional analysis has also been conducted for day-time data and night-time data respectively. We observed that detection accuracy is lower during nighttime. An average model might not be adequate for detections both in daytime and nighttime even with system parameter adjustments. This suggests that a suite of models might be needed for varying lighting conditions.

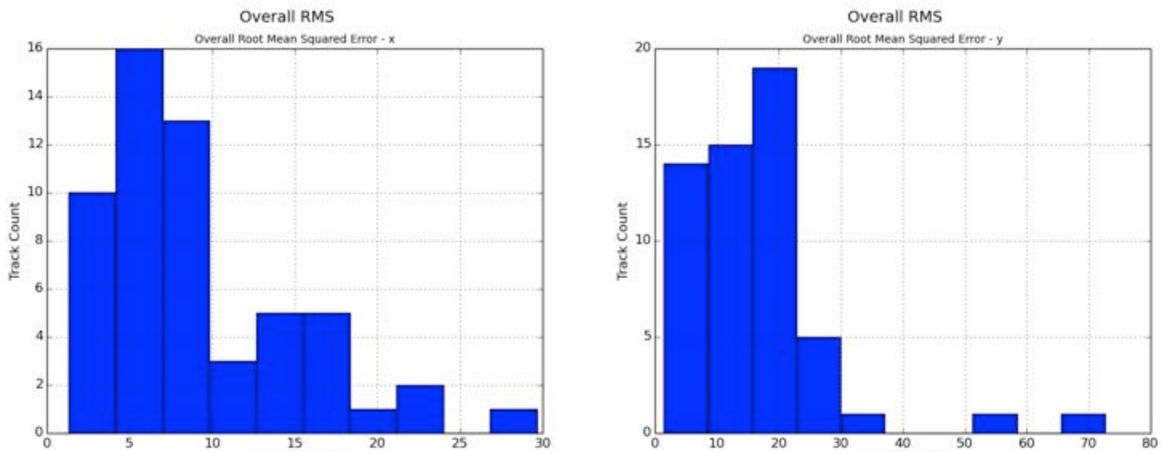

Figure 16. Histogram of x-axis and y-axis RMS Error in Pixels over All Tracks

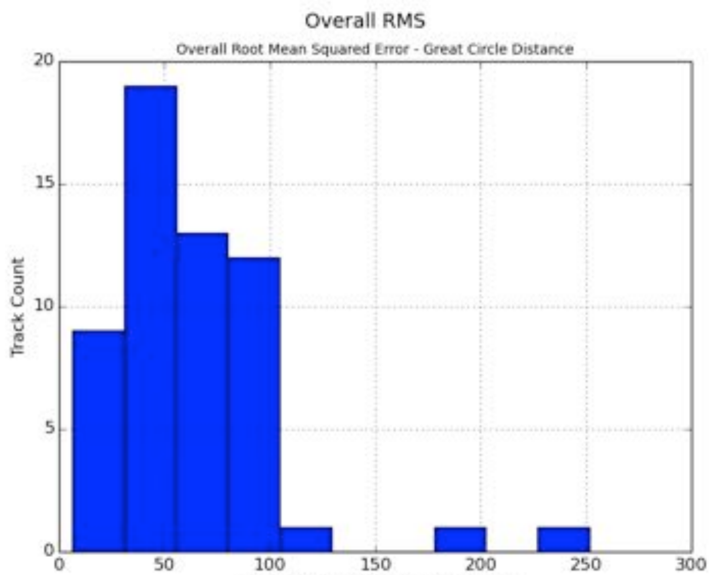

Figure 17. Histogram of RMS Error in Feet over All Tracks

Table 2 summarizes the tracking dropout rate for different weather and lighting conditions. Overall speaking, tracking dropout rate is lower during the day and higher during the night and the overall average dropout rate is 1.8%. Since the aircraft detection model was trained using data from 6 am to 9 pm, in which the amount of daytime data is larger than the night-time data, it is conceivable that aircraft detection would perform better during the day. Nonetheless, the small overall dropout rate is encouraging.

Table 2. Tracking Dropout Percentage over Different Weather/Lighting Scenarios

Condition	Clear Day	Clear Night	Overcast Day	Overcast Night	Rain Day	Rain Night	Snow Day	Snow Night
%	0.38	2.8	1.9	9.5	1.6	3.3	0	1.1

V. Concluding Remarks

The objective of the current research is to develop a Vision-BAsed Surveillance System (VBASS) for ramp areas in airports. Vision-based surveillance technologies are based on surveillance cameras located at certain vantage points in the ramp area. VBASS can be used in real-time mode or it can be used in post-processing (playback) mode. In real-time mode VBASS can help monitor ramp-area operations in an automated manner; it can also aid planners such as the Spot and Runway Departure Advisor (SARDA) by providing real-time updates on the pushback status of the aircraft. In post-processing mode VBASS can help identify from video surveillance data, patterns of interest in ramp area operations such as the paths from the gates to spot, and the transit time from the gate to the ramp spot.

Under this research effort, we installed seven cameras at the Charlotte (CLT) Douglas International Airport and demonstrated that VBASS is capable of tracking aircrafts under different lighting conditions using the streaming videos replayed from the video recordings obtained from the CLT cameras. We also demonstrated that the standalone VBASS is capable of taking in the real-time video data feed and outputting the surveillance results to a top-down display at NASA's CLT lab. Future work includes enhancing the robustness of the VBASS system as well as developing software tools to automate the airport adaptation capability of the VBASS system.

Acknowledgments

This development effort was performed under support from NASA Contract Number NNX15CA65C. The authors would like to thank our technical monitor Dr. Yoon Jung of NASA Ames Research Center for his tremendous support throughout the project. We would like to thank ATD-2 team lead Shawn Engelland for overseeing the project and for providing technical support that bridges VBASS and ATD-2 development. Thanks are also due to numerous ATD-2 team developers at Ames who provided support for OSI CLT server installation, CLT camera connectivity, and CLT lab network access, including Kevin Day, J. D. Sanders, and Shivanjli Sharma. We also would like to thank Dr. Miwa Hayashi for her insightful suggestions throughout the VBASS development.

References

[1] Concept of Operations for the Next Generation Air Transportation System, v. 2.0, Joint Planning and Development Office, June 13, 2007.

[2] Swenson, H., Barhydt, R., Landis, M., Next Generation Air Transportation System (NGATS) Air Traffic Management (ATM) – Airspace Project, NASA Ames and Langley Research Centers, June 1, 2006.

[3] Hinton, D., J. Koelling, and M. Madson, Next Generation Air Transportation System (NGATS) Air Traffic Management (ATM) – Airportal Project, Reference Material, External Release Version, NASA, May 23, 2007.

[4] "FAA's NextGen Implementation Plan," Federal Aviation Administration, Washington, DC, June 2008.

[5] "FAA's NextGen Implementation Plan 2009," Federal Aviation Administration, Washington, DC, revised on February 10, 2009.

[6] "Airport Rules and Regulations," Port Authority of New York and New Jersey, Aviation Department, August, 2009.

[7] Guilloton, A., Arenthens, J-P., Escher, A-C., Macabiau, C., and Koenig, D., "Multipath Study on the Airport Surface," IEEE 2012.

[8] http://autoscope.com/products/solo_terra.htm, as viewed on August 9, 2012.

[9] http://www.lumenera.com/products/surveillance-cameras/index.php, as viewed on August 9, 2012.

[10] http://www.sensorsinc.com/, as viewed on August 9, 2012.

[11] http://www.flir.com/cvs/americas/en/security/view/?id=44284, as viewed on August 9, 2012.

[12] Jung, Y. C., Hoang, T., Montoya, J., Gupta, G., Malik, W., and Toibas, L., "A Concept and Implementation of Optimized Operations of Airport Surface Traffic," Proceedings of the 2010 AIAA Aviation, Technology, Integration and Operations Conference, September, 2010.

[13] System Wide Information Management, https://www.faa.gov/nextgen/programs/swim/.

[14] P. Felzenszwalb, D. McAllester, D. Ramaman, "A Discriminatively Trained, Multiscale, Deformable Part Model," *Proceedings of the IEEE CVPR*, 2008.

[15] P. Felzenszwalb, R. Girshick, D. McAllester, D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 32, No. 9, September 2010.